



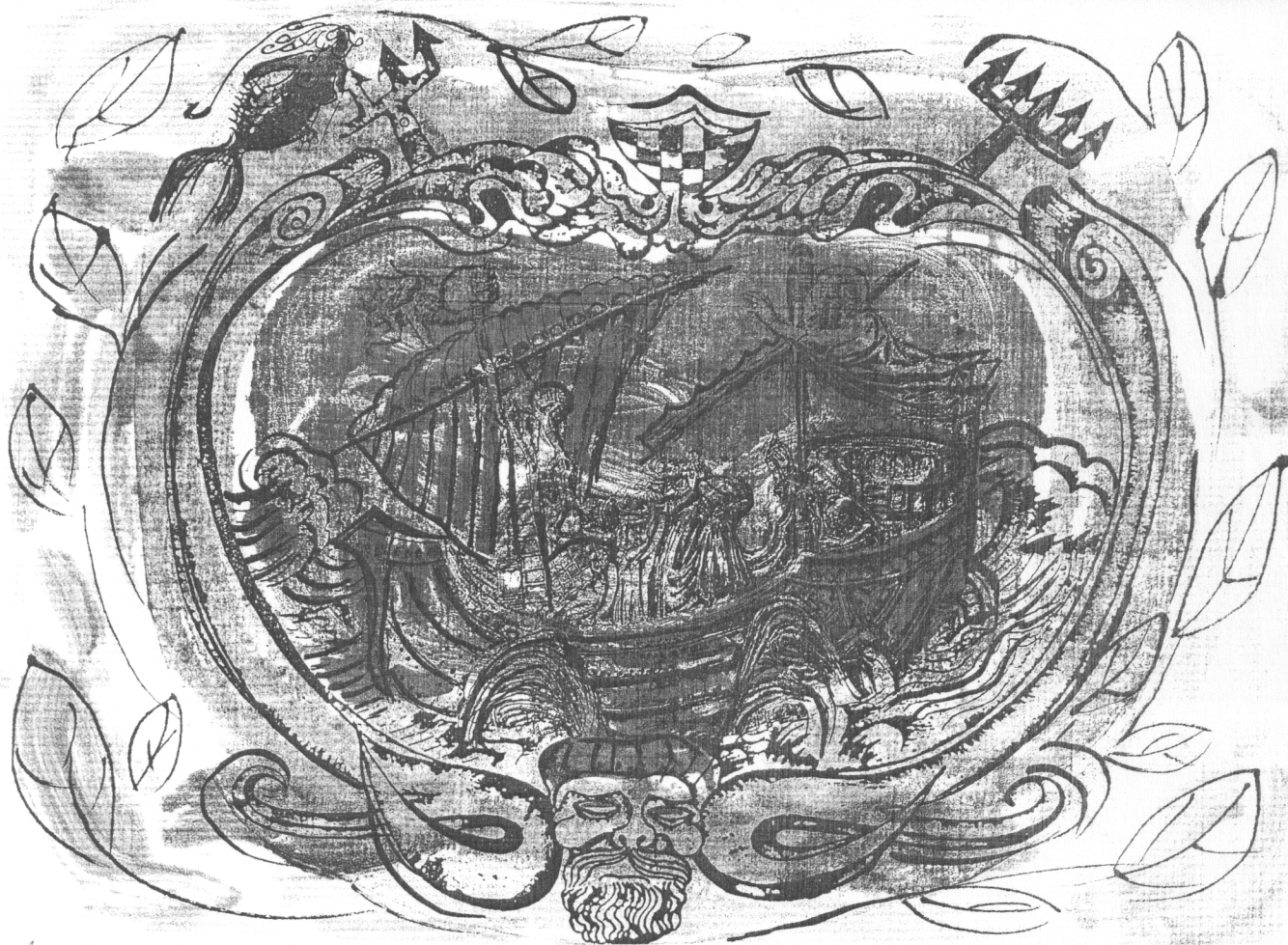
Elmar



IEEE
Croatia Section

PROCEEDINGS VIPromCom '99

1st ELMAR International Workshop on
Video Processing and Multimedia Communications



1998.

h. Muiš

Edited by:
Mislav Grgić

ZADAR • Croatia
June 23 to 25. 1999

Cataloging-in-Publication Data

CIP - Katalogizacija u publikaciji
Nacionalna i sveučilišna knjižnica, Zagreb

UDK 621.39 (063)

ELMAR International Workshop on Video
Processing and Multimedia Communications
(1 ; 1999; Zadar)

Proceedings VIPromCom'99 / 1st Elmar
International Workshop on Video Processing
and Multimedia Communications, June 23-25.
1999 Zadar, Croatia ; edited by Mislav Grgić.
- Zadar : Croatian Society Electronics in Marine,
1999. - 313 str. : ilustr. ; 30 cm

Bibliografija uz svaki rad. - Kazalo.

ISBN 953-96030-5-6

1. Grgić, Mislav

990601029

Publisher: Croatian Society Electronics in Marine - ELMAR, Zadar
Editor: Mislav Grgić
Front Cover: Sepulchre of St. Šimun in Zadar, detail; interpreted by artist
Mrs. Ljerka Njerš, from the opus "Hommage á Croatian Cultural Heritage"

DTP and Printed by: BS - ZAGREB

ISBN: 953-96030-5-6

© by Croatian Society Electronics in Marine - ELMAR, Zadar, 1999.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission from the Croatian Society Electronics in Marine - ELMAR, Zadar.

All papers were reproduced exactly as received, on author's responsibility.

Printed in Croatia.

| | |
|---|-----|
| R. Šapina; CROATIA: Quality of Edge Detection Image Analyses | 125 |
| E. Diaconescu, M. Calugareanu; ROMANIA: Fuzzy Regional Method for Image Segmentation | 131 |
| M. Oravec, J. Polec; SLOVAKIA: Image Interpolation by Discrete Orthogonal Transforms and Multilayer Perceptron | 135 |
| H. Balaško, A. Luštica; AUSTRIA / CROATIA: Advantages of DSP Camera Processing | 145 |
| P. Skočir, B. Marušič, J. Tasič; SLOVENIA: A Postprocessing Approach to Wavelet Ringing Removal | 155 |
| K. Kerš, S. Grgić, M. Grgić; CROATIA: Two-Dimensional Wavelet Transform in Image Compression | 163 |
| M. Ravnjak, M. Grgić, B. Zovko-Cihlar; CROATIA: Features of Wavelet Filters in Still Image Compression | 177 |
| N. Stevanovic, M. Hillebrand, B.J. Hosticka, J.E. Santos Conde, A. Teuner; GERMANY: High Speed Photography, 3D Optical Measurement, and Intelligent Surveillance Applications Using CMOS Imager | 187 |

SESSION: MULTIMEDIA COMMUNICATIONS

| | |
|---|-----|
| J. Pavasović, K. Skala, J.F. Tasič; SLOVENIA / CROATIA: Using Agents for Engaged Learning | 203 |
| V. Kedmenec, M. Gerenčir, M. Grgić, B. Zovko-Cihlar; CROATIA: Utilizing MPEG Compression for High Quality Videoconference | 211 |
| V. Klepac; CROATIA: Digital Tape Formats, Compression and Networking | 217 |
| M. Krčum, I. Vujović, I. Kuzmanić; CROATIA: Multimedia Communication - Wavelet Analyses | 223 |
| P. Košut, J. Polec; SLOVAKIA: Generalized Combined ARQ Scheme for Data Transmission | 237 |
| M. Suknaić, M. Grgić, B. Zovko-Cihlar; CROATIA: The Design of 3rd Generation Air Interface | 241 |
| T. Kos, B. Zovko-Cihlar, S. Grgić; CROATIA: Nonlinear Distortion in Broadband CATV Amplifiers | 255 |
| S. Rozman, F. Solina; SLOVENIA: Image Database Queries Based on Interest Points | 263 |

IMAGE DATABASE QUERIES BASED ON INTEREST POINTS

Stanislav Rozman

Franc Solina

Computer Vision Laboratory
Faculty of Computer and Information Science

University of Ljubljana

Tržaška 25, SI - 1000 Ljubljana, Slovenija

Stanislav.Rozman@mail.intramit.si

franc@fri.uni-lj.si

The method described uses Harris - Stephens's combined corner and edge detector and thresholding to determine interest points. A set of positions of interest points in an image, stored in rounded, normalized polar coordinates is used as a feature of an image in image database. Distance function is based on a number interest points with same positions in two images. The method is acceptable in terms of speed and can to some extent retrieve images with similar contours independently of their color information.

Keywords: image queries, image indexing, interest points, corner and edge detection

1. INTRODUCTION

Development in computer technology, increasing popularity of the Internet and growing use of multimedial documents result in a growing demand for multimedia management algorithms. One of the approaches that help handling large image databases is *content based image retrieval (CBIR)*. Such systems extract certain image properties, using the acquired image feature as an index and express image similarity by calculating the distance between features. Many existing applications perform very well, especially using color and texture information of images, with demonstrations available over the internet [1, 2, 3, 4]. Current development is discussed in [5]. Especially image queries based on color information proved very effective. However, images retrieved via their color information are not always to ones expectations or satisfaction.

This paper examines possibilities of a method that would extract similar images independently of their color information using interest points to mark image features or contours (i.e. edges, corners) with performance fast enough to be useful for an application purpose. Harris and Stephens combined corner and edge detector was used in [6, 7] to determine interest points in order to identify images in a large database using a rotated, scaled or partly visible input image. The main question was whether the repeatable results of Harris and Stephens detector are useful in retrieving similar images as well. This detector was used to determine candidates for interest points and then thresholding and reduction were used to obtain a smaller group of points that would give as much information about an image as possible. The positions of these points were transformed into rounded and normalized polar coordinates, giving an universal representation for each image. Set of such coordinates was used as an image feature and the distance function between two images was based on number of interest points with same positions in both images.

After initial testing on a smaller set of images to determine the ideal parameter setting a test on a large image database that was previously used by algorithms based on color information [8, 9] was run. The performance shown was satisfying in terms of speed and an encouraging number of similar images were retrieved. Retrieved images were also observed in terms of context, color, size and orientation independency.

2. INTEREST POINTS

Interest points are local features at which the signal of the input image changes two-dimensionally. Harris and Stephens detector was used because it is known to give most repeatable results under varying conditions. In [10] it is described as follows:

$$R(x, y) = \det \hat{C} - 0.04 \cdot \text{tr}^2 \hat{C} \quad (1)$$

where \hat{C} is defined as:

$$\begin{bmatrix} I_x^2 & I_x \cdot I_y \\ I_x \cdot I_y & I_y^2 \end{bmatrix} \otimes e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2)$$

and I_x, I_y are the first derivatives of image signal.

The values of $R(x, y)$ are thresholded to obtain a minimum set of points with as much image descriptive information as possible.

3. SET OF POINTS AS AN IMAGE FEATURE

3.1. Reduction of candidates

Once candidates for interest points using Harris-Stephens detector and thresholding are determined, further selection is made. First, neighboring candidates are deleted to reduce the number of points since we have no need for points being joined in lines or even regions. Actually, our intention is the opposite. Furthermore, points lying close to any of the image boundaries (less than 5%) are ignored as well. This was necessary because several images had frames, signatures or text which all severely disrupted the process of selection of interest points. A five percent margin chosen arbitrarily compensated most of these side-effects with least loss of desired features.

3.2. Determining the image "centre"

In order to transform interest points' positions into polar coordinates, first the centre of the image has to be determined. The work of Demšar and Solina [11] suggests that features derived from the central part of the image are decisive. The idea of this method, however, is that the most informational features of the image are usually positioned in the central area of the image, but not necessarily exactly in the centre. Since centre location has a very important effect on image feature, an as-exact-as-possible adjustment is necessary. The locations of interest points determined can be used to estimate the "logical" centre. In this procedure only the interest points that are at least 20 % away from any boundary, are used, since we don't want side areas to have too much effect. At this point, resolution is temporarily decreased by 10 times. This is used to discriminate areas with high density of interest points. On the remaining points median of x and y coordinates is used to determine the centre of the image.

3.3. Transforming to polar coordinates

Rounded polar coordinates are used to decrease resolution and to discriminate significance of points away from the centre. The further away from the centre, the less and less points have to be used to cover "dense" areas of interest points, and the location of the points is less exact, giving more tolerance for matching, as well. The values of radius are normalized using the length of the diagonal of the image thus achieving a standard scale for all images. Units proven best in preliminary tests were 180 units for angle (2 degrees) and 150 pixels for radius.

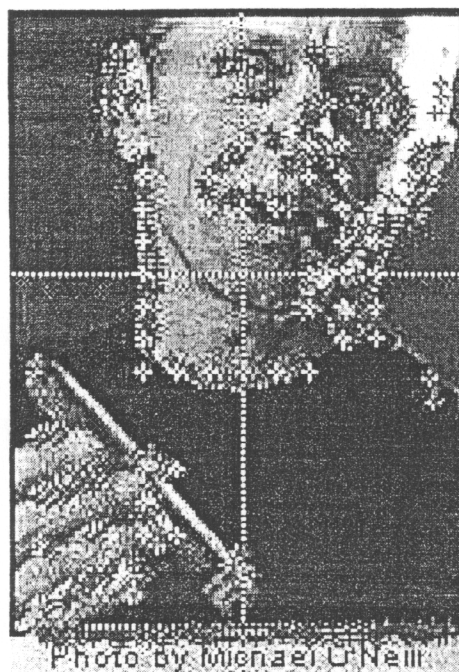


Figure 1: A sample of an output image, created by the image feature calculation program. Small crosses of different colours indicate candidates for interest points, actual interest points in polar coordinates and "heavy" points that help determine the centre. The centre itself is shown by a dotted cross through the whole image. Such an output image helps the designer while setting the parameters as well as assists in results evaluation since it helps to explain system behavior.

4. DISTANCE FUNCTION

As previously mentioned, the main idea of the distance function is to show similarity between two images based on the number of interest points with same positions in two images. Finally, the following form of the distance function was used:

$$D_{ij} = \begin{cases} \infty & : m_{ij} = 0 \\ \frac{N_i \cdot N_j}{m_{ij}} - 1 & ; \text{otherwise} \end{cases} \quad (3)$$

where:

D_{ij} means distance between images S_i and S_j ,

N_i means number of interest points in image S_i ,

m_{ij} means number of interest with same positions in images S_i in S_j .

Obviously, $D_{ij} \in [0, \infty)$, $D_{ii} = 0$, $D_{ij} = D_{ji}$.

5. PERFORMANCE

A Pentium personal computer with 166 MHz and 64 MB RAM was used in final tests on a database with 1164 images. Programs for calculating image feature and distance function were developed in C programming language. The image feature calculation program needed 12052 [sec]. The program which calculated distance from each to each image then giving best

seven and worst three similarities, needed 14594 [sec]. Therefore, an application which would compare a new image to an existing database of 1200 images would take an average of 10 [sec] to calculate the feature for the new image and further 25 [sec] for evaluating the distance to all the images in the database. Feature calculation program was tested on a Pentium II with 233 MHz and 128 MB RAM and the execution times were shorter by a factor of three! Therefore, performance can be evaluated as satisfying, perhaps with desire to improve the distance function, since it grows linearly with the size of database.

6. RESULTS

Evaluation of results is a tricky task because of its subjectivity. For 238 images in the database, their seven best neighbors (smallest distances) and three worst (greatest distances) were previewed and evaluated for similarity and dissimilarity in terms of human perception as well as in terms of expected feature performance. The following characteristics were observed:

- similarity of first image to the reference image
- similarity of first seven images to the reference image
- dissimilarity of last three images to the reference image
- context classification: how (if) the method joined images of similar context
- color independency: how well (if) method retrieved similar images with little similarity in color information
- independency on size and orientation

6.1. Evaluation

The set consisting of an reference image and its neighbors (7 + 3) was given a grade according to this scale:

| grade | similarity | | |
|-------|-------------|---|------------|
| | first image | next six | last three |
| R | yes | a lot | none |
| Q | yes | a lot | some |
| S | yes | some | none |
| P | no | a lot | none |
| O | no | a lot | some |
| M | no | promising, but not enough some similarity, but very questionable | |
| A | no | | |
| N | no | none | none |
| F | no | none | some |

Table 1: Grading scale

6.2. Examples

Three examples are shown in figures 2, 3 and 4: the reference image is first, and then the seven similar images are displayed in rows from left to right with distance values indicated below each image.

| group | grade | no. | [%] | total | |
|-----------|-------|-----|-------|-------|---------|
| success | R | 23 | 9.66 | 46 | 19.32 % |
| | Q | 11 | 4.62 | | |
| | S | 12 | 5.04 | 63 | 26.46 % |
| | P | 17 | 7.14 | | |
| promising | O | 14 | 5.88 | 70 | 29.41 % |
| | M | 26 | 10.92 | | |
| | A | 30 | 12.61 | | |
| failure | N | 77 | 32.36 | 105 | 44.13 % |
| | F | 28 | 11.77 | | |

Table 2: Evaluation results for 238 images and their neighbors

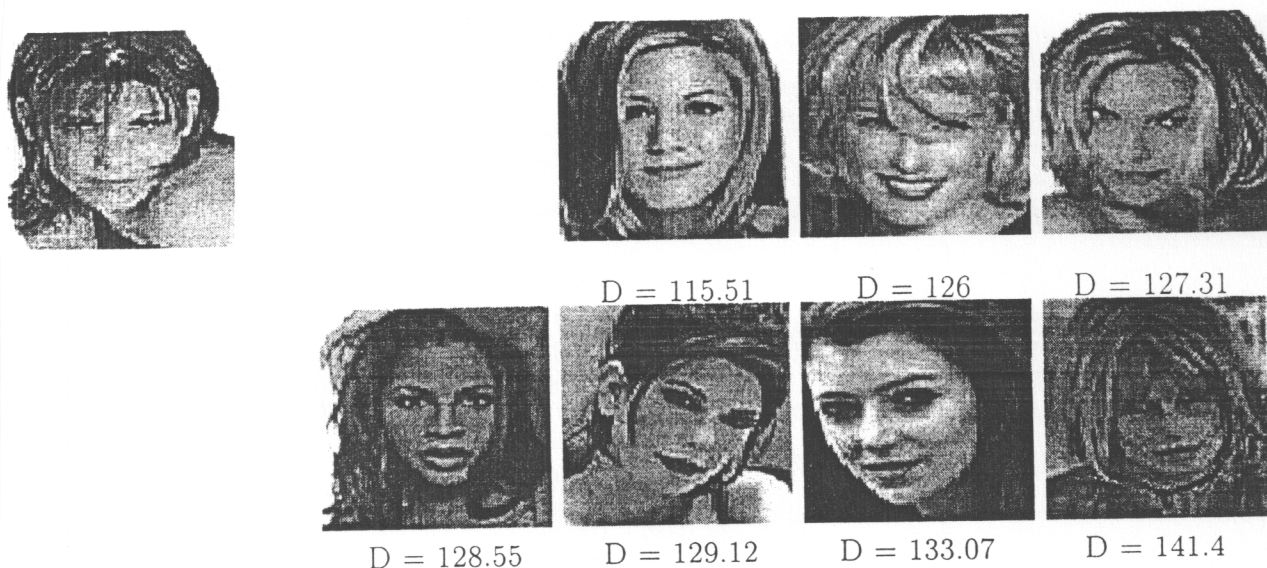
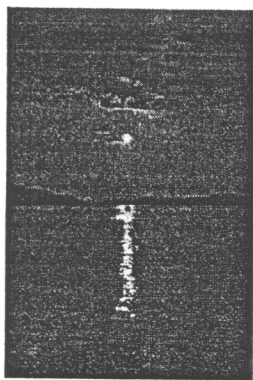


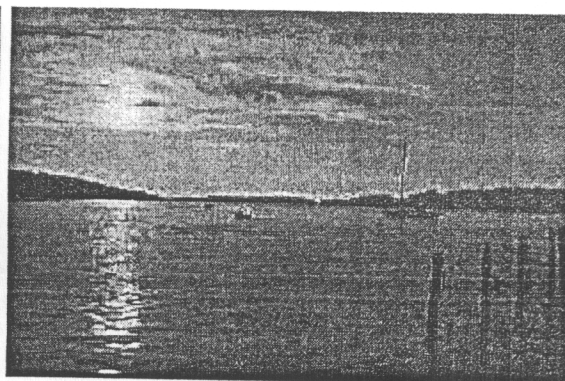
Figure 2: Girls' faces: Different hair color and skin tan, slightly different orientation. Graded: R

6.3. Findings

The method is to some extent successful in separating the image database into subsets by image context. Retrieved similar images very often have different color information and would therefore be irretrievable with a color-based method. Previewed and evaluated examples included retrieval of images that were a part of original - reference image and scaled images. This is expected considering the results of [6, 7]. However, the evaluation has also shown that the value of the distance function and results derived using it depend heavily on the number of interest points in an image. The greater the number, the smaller the distances since probability of a greater number of matches rises. Dealing with this side effect could further improve the results.



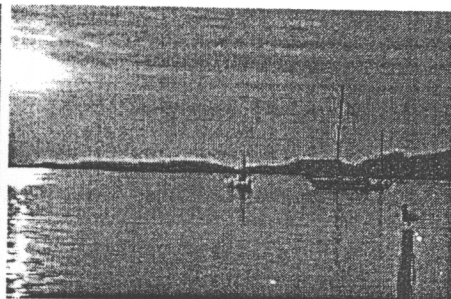
$D = 62.87$



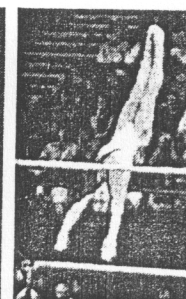
$D = 68.11$



$D = 71.28$



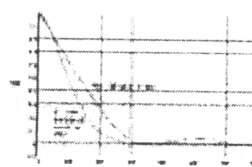
$D = 81.81$



$D = 87.78$

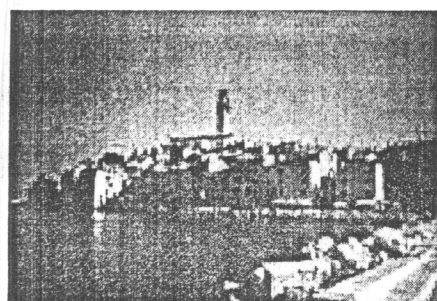


$D = 87.88$

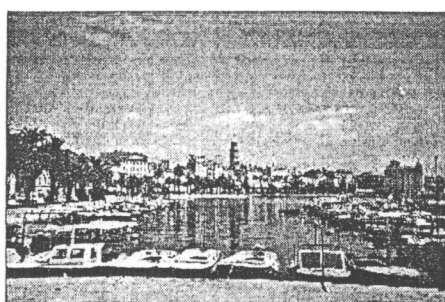


$D = 100.08$

Figure 3: Bays: Different locations, entirely different colors. A gymnast's image and a diagram interfere. Graded: **R**



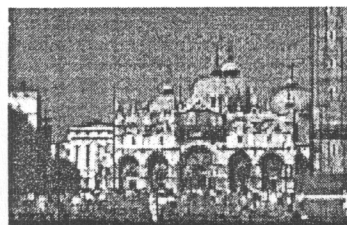
$D = 16.76$



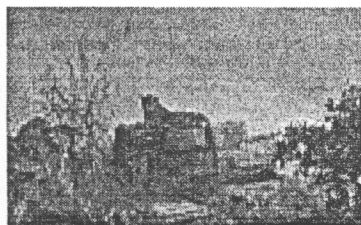
$D = 18.28$



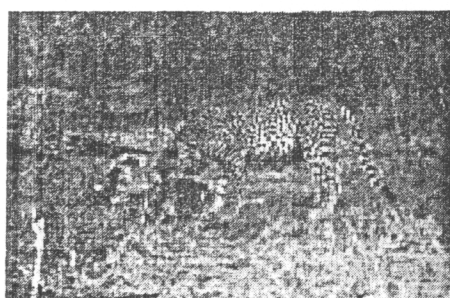
$D = 18.95$



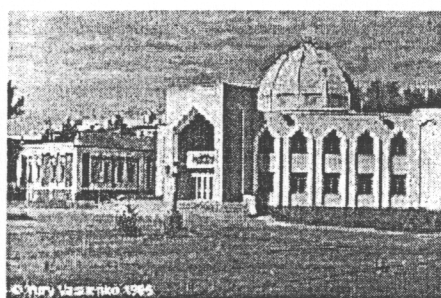
$D = 19.14$



$D = 19.43$



$D = 21.95$



$D = 22.02$

Figure 4: View of Rovinj. Method successfully retrieves an image of Split and the San Marco square plus another urban scene, the "safari" scenes, however, interfere. Graded: P

7. CONCLUSION

This paper proves Harris and Stephens corner and edge detector to be very useful for querying similar images by interest points matching. Using interest points' location in polar coordinates also proved suitable. Main effort in further development should be put in modifying the distance function.

Naturally, the method was not expected to outperform the methods based on color information. However, it was able to retrieve similar images to some extent, and has fully met the expectations as color independence is considered.

With further development this method could be useful as a foundation for a contour-based image retrieval application. Even more prospects can be seen if used in combination with other types of methods (color, texture).

REFERENCES

- [1] QBIC.
<http://www.qbic.almaden.ibm.com/>
- [2] VIRAGE.
<http://www.virage.com/virdemo.html>
- [3] MetaSEEk.
<http://www.ctr.columbia.edu/metaseek/>
- [4] BSQ.
<http://diana.fri.uni-lj.si/bsq/>
- [5] H.S. Stone, Image Libraries and the Internet, *IEEE Communications magazine*, January 1999, vol. 37, no. 1 pp. 99 - 106
- [6] C. Schmid and R. Mohr, Local grayvalue invariants for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, May 1997 - vol.19, no. 5 pp. 530 - 535
- [7] J.M. Marie-Julie and H. Essafi, Using Ifs and moments to build a quasi invariant image index. *Proceedings of 5th European Conference on Computer Vision (ECCV'98)*, Freiburg, Germany, June 1998, Volume I, pp- 578 - 594
- [8] D. Radolović, *Image database queries based on color information*, BS thesis, University of Ljubljana, June 1998
- [9] J. Demšar and F. Solina, Using machine learning for content-based image retrieving, *Proceedings of the 13th International Conference on Pattern Recognition*, International Association for Pattern Recognition, IEEE Computer Society Press, Vienna, Austria, August 1996, vol. IV, pp. 138-142,
- [10] R. Deriche and G. Giraudon, A computational approach for corner and vertex detection, *International Journal of Computer Vision*, vol. 10, no. 2 pp. 101-124
- [11] J. Demšar and F. Solina, Searching for faces in image data bases using machine learning. *Speech and image understanding, IEEE Slovenia section*, Ljubljana, 1996, pp. 199-207.
- [12] E. Truco and A. Verri, *Introductory Techniques for 3-D Computer Vision*, Prentice-Hall, 1998

- [13] R. Jain, Visual Information Management, *Communications (of the ACM)*, December 1997 - vol. 40, no. 12 pp. 31 - 32
- [14] A. Gupta, S. Santini and R. Jain, In search of Information in visual media, *Communications (of the ACM)*, December 1997 - vol. 40, no. 12 pp. 35 - 42
- [15] S. Rozman, *Image database queries based on interest points*, BS thesis, University of Ljubljana, February 1999